

# **Assessment and Evaluation of Individually Calibrated Journey-To-Crime Geographic Profiling Models**

**Michael Leitner**

Department of Geography and Anthropology  
Louisiana State University

**Tania Pal**

Department of Geography and Anthropology  
Louisiana State University



# Research Question 1

- Comparing the accuracy of Journey-To-Crime (JTC) Geographic Profiles (GP) created from
  - individually calibrated distance decay functions with
  - using the default values in CrimeStat (3.0)



## Research Question 2

- Comparing the accuracy of individually calibrated JTC GP with alternative modern GP models (Rigel and Dragnet) and simple spatial distribution measures (spatial mean, spatial median, center of minimum distance).
- This research question is answered by comparing the results from this study with results from previous research.



# Rationale

- If there were no difference, whether JTC GP are created from default or individually calibrated distance decay functions, then
  - Default parameter values should be used when creating JTC GP
  - Distance decay functions do not need to be individually calibrated
- This would save time and resources (personal, money)
- This comparative analysis has never been done before.



## Definition - Geographic Profile

... is a decision support tool used by law enforcement to make estimates about the likely location of a serial offender's haven.



## Comparison - Geographic Profiling Models

Paulsen, Derek J. 2006 “Connecting the Dots: Assessing the Relative Accuracy of Geographic Profiling Software”. *Policing: An International Journal of Police Strategies and Management*. Vol 29, Issue 2, pp. 306-334.

Compares various accuracy measures across different GPs

- JTC GP (using *CrimeStat III* default parameters)
- Rigel
- Dragnet
- simple spatial distribution measures (spatial mean, spatial median, center of minimum distance)



## Results - GP Comparison (Paulsen, 2006)

- Probability strategies (JTC GP, Rigel, Dragnet) are not substantially more accurate than spatial distribution measures (spatial mean, spatial median, center of minimum distance)
- Consistent with previous findings (Levine 2002, Snook *et al.* 2005)



## Research - GP Comparison

Extend the research by Paulsen (2006) by using individually calibrated distance decay functions instead of the default values in CrimeStat (3.0).

- Same data set
- Same size of Geographic Profile (x-, y-coordinates of lower left and upper right corner of GP)

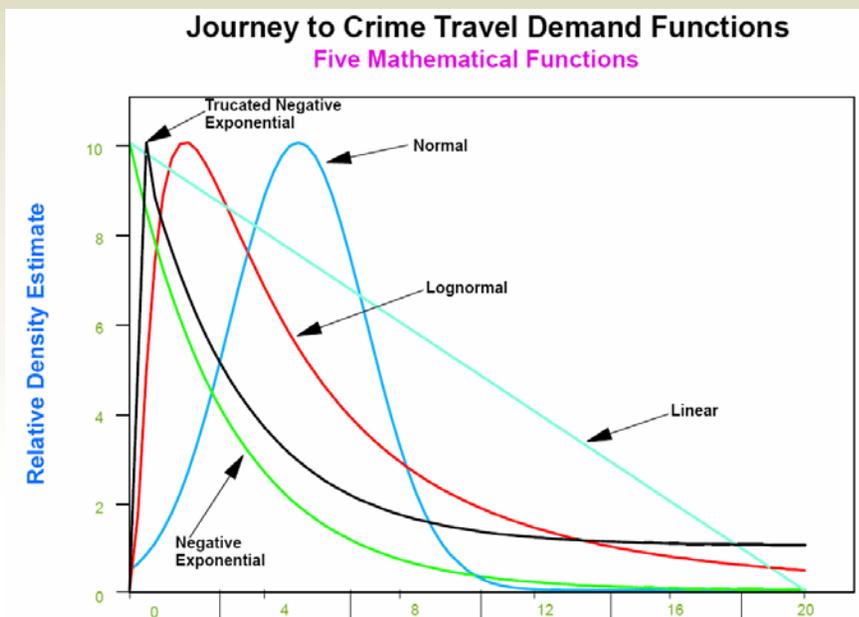


## Data & Study Area

- 247 Serial Crimes
- Nine different crime types
  - Larceny (51 serial crimes)
  - Arson (4 serial crimes)
  - Auto theft (31 serial crimes)
  - Robbery (commercial-76, street-17, mixed-15)
  - Rape (1 serial crime)
  - Burglary (residential-51, commercial-1)
- Both crime locations and actual “haven” known
- Three or more offenses in each crime series
- 1994 – 1997
- Baltimore County, Maryland

# JTC GP Method

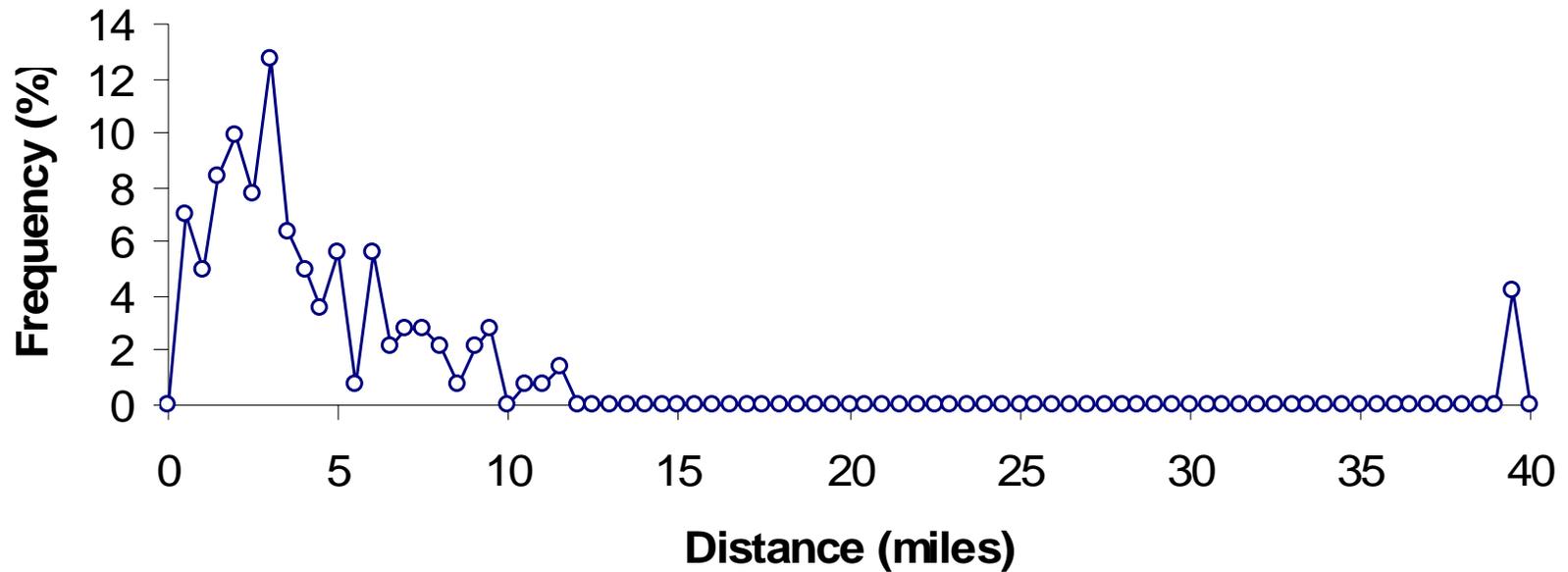
- **Calibration group:** Many (serial) offenders for which travel patterns to and from the crime location are known
- These travel patterns are modeled with various **distance decay functions** (modeling=estimation of parameters, calibration)
- **Test group:** One serial offender with known crime locations
- **JTC GP:** It integrates the crime location from the test group with the calibrated distance decay function.





# Auto Theft Crime Scene Distribution

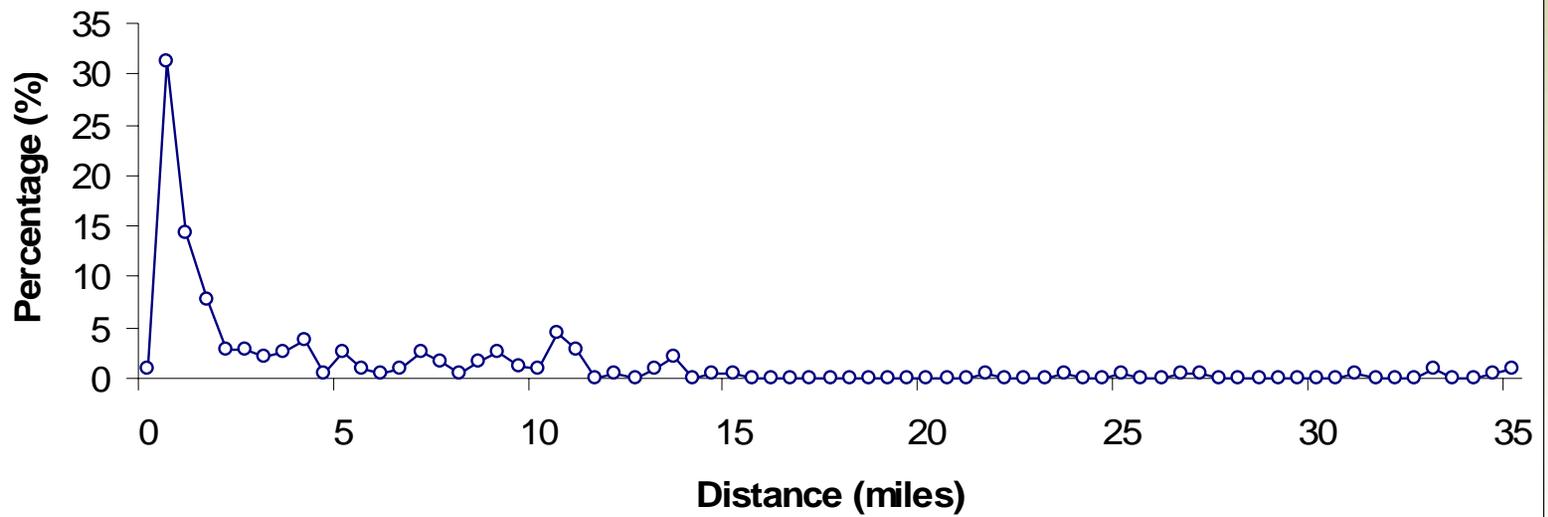
0.50 Mile Bin Distance Interval  
Baltimore County Serial Auto Theft 1994 - 1997





## Residential Burglary Crime Scene Distribution

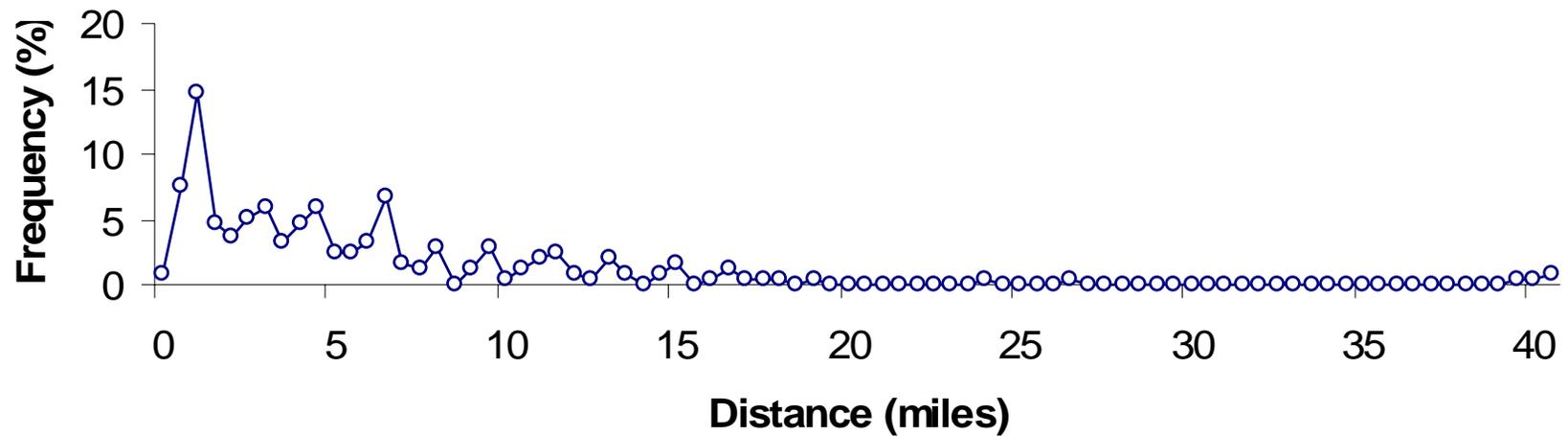
0.50 Mile Bin Distance Interval  
Baltimore County Serial Residential Burglary 1994 - 1997





## Larceny Crime Scene Distribution

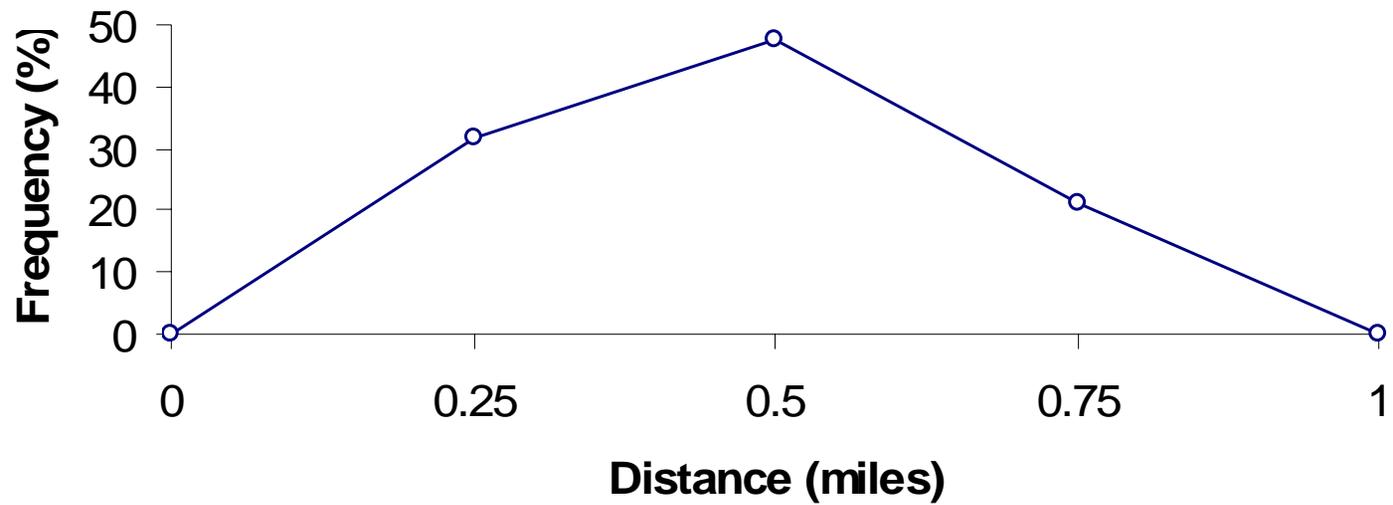
0.50 Mile Bin Distance Interval  
Baltimore County Serial Larceny 1994 - 1997





## Arson Crime Scene Distribution

0.25 Mile Bin Distance Interval  
Baltimore County Serial Arson 1994 - 1997

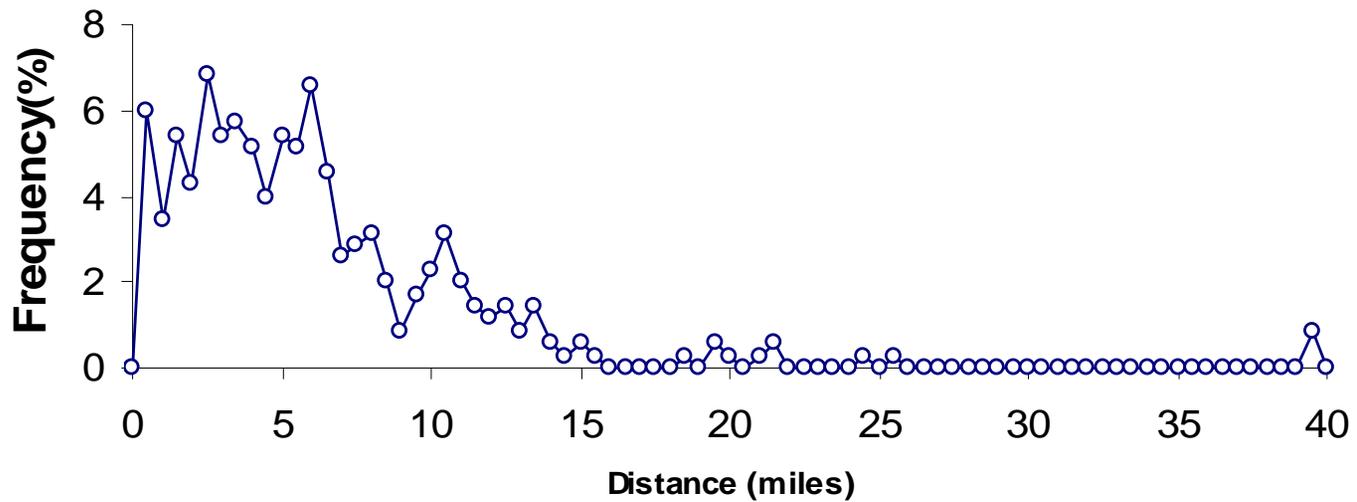




# Commercial Robbery Crime Scene Distribution

0.50 Mile Bin Distance Interval

Baltimore County Serial Commerical Robbery 1994 - 1997

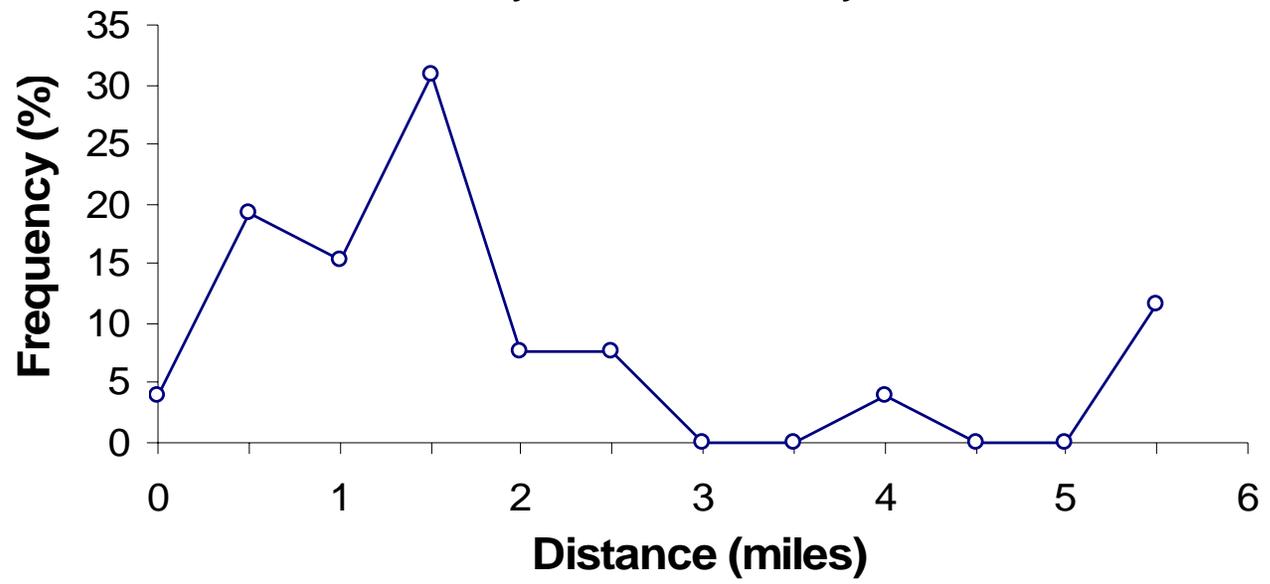




# Mixed Robbery Crime Scene Distribution

0.50 Mile Bin Distance Interval

Baltimore County Serial Mixed Robbery 1994 - 1997

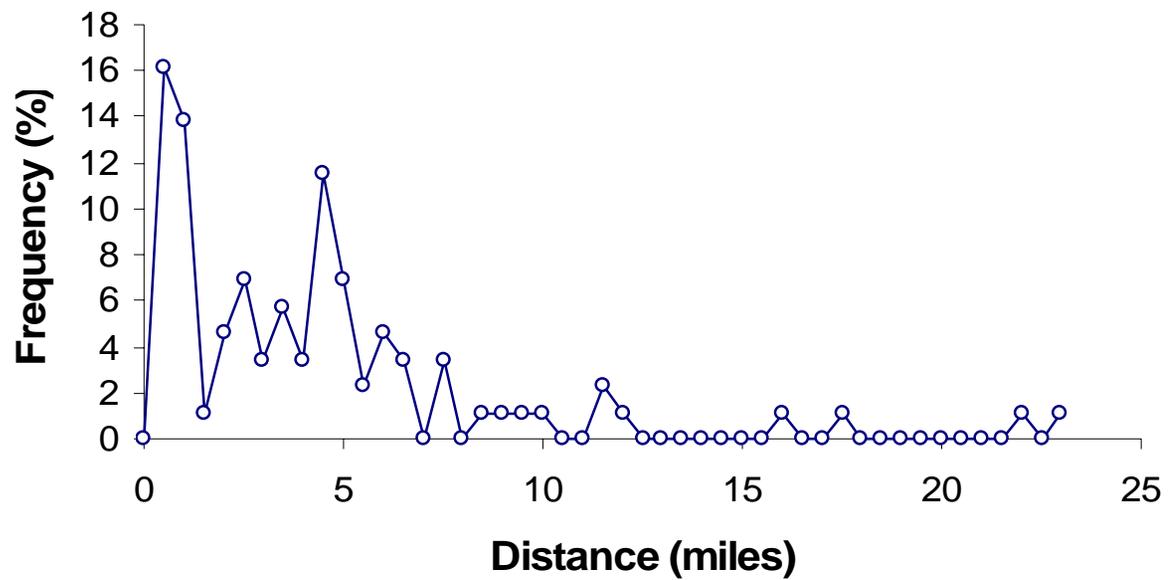




## Street Robbery Crime Scene Distribution

0.50 Mile Bin Distance Interval

Baltimore County Serial Street Robbery 1994 - 1997





## Procedure - Using default parameters for calibrating distance decay functions

- The procedure is explained using the larceny dataset with 51 serial crimes
- **Step 1:** Select the first serial crime from the larceny dataset (**test group**).
- **Step 2:** Create five different JTC GP using the default parameter values for each of the five different distance decay functions implemented in CrimeStat (3.0).
- **Step 3:** Repeat Steps 1-2 for each serial crime in the larceny dataset.

This procedure results in 51 JTC GP for each of the five default calibrated distance decay functions.



## Procedure - Using parameter values from individually calibrated distance decay functions

- The procedure is explained using the larceny dataset with 51 serial crimes
- **Step 1:** Remove the first serial crime (**test group**) from the larceny dataset.
- **Step 2:** Calibrate five different distance decay functions from the remaining 50 serial crimes (**calibration group**).
- **Step 3:** Create a JTC GP for each of the five different distance decay functions for the first larceny serial crime removed from the dataset in Step 1.
- **Step 4:** Repeat Steps 1-3 for each serial crime in the larceny dataset.

This procedure results in 51 JTC GP for each of the five individually calibrated distance decay functions.



## Comparison between GP calculated from default and individually calibrated distance decay functions

- The comparison is conducted for
  - All crime types
  - Five different distance decay functions (linear, negative exponential, truncated negative exponential, normal, lognormal), and
  - Three different comparison measurements (error distance, search area size, and hit score percentage).
- For each crime type, distance decay function, and comparison measurement a paired-samples t-Test is used to compare the individually calibrated with the default calibrated JTC GP.



# JTC GP in CrimeStat III

The screenshot shows the CrimeStat III software interface. The main window title is "CrimeStat III". The "Options" tab is selected, and within it, the "Crime travel demand" sub-tab is active. Below the sub-tab are four smaller tabs: "Primary File", "Secondary File", "Reference File", and "Measurement Parameters".

The "Primary File" tab is selected, showing a file path: "D:\Students\PaTania\Tania's Thesis\Baltimore247Crime Series\Burglary\ResidentBur". A "Select Files" button is to the right, and "Edit" and "Remove" buttons are below the path.

The "Variables" section contains a table with the following columns: Name, File, Column, and Missing values.

| Variables Name | File  | Column | Missing values |
|----------------|---|--------|----------------|
| X              | D:\Students\PaTania\Tania's Thesis\Baltimore247Crime Series\{ | INCIDX | <Blank>        |
| Y              | D:\Students\PaTania\Tania's Thesis\Baltimore247Crime Series\{ | INCIDY | <Blank>        |
| Z (Intensity)  | D:\Students\PaTania\Tania's Thesis\Baltimore247Crime Series\{ | <None> | <Blank>        |
| Weight         | D:\Students\PaTania\Tania's Thesis\Baltimore247Crime Series\{ | <None> | <Blank>        |
| Time           | D:\Students\PaTania\Tania's Thesis\Baltimore247Crime Series\{ | <None> | <Blank>        |
| Directional    | D:\Students\PaTania\Tania's Thesis\Baltimore247Crime Series\{ | <None> | <Blank>        |
| Distance       | D:\Students\PaTania\Tania's Thesis\Baltimore247Crime Series\{ | <None> | <Blank>        |

Below the table are three sections of options:

- Type of coordinate system:**  Longitude, latitude (spherical),  Projected (Euclidean),  Directions (angles)
- Data units:**  Decimal Degrees,  Feet,  Meters,  Miles,  Kilometers,  Nautical miles
- Time Unit:**  Hours,  Weeks,  Days,  Months,  Years

At the bottom of the window are three buttons: "Compute", "Quit", and "Help".

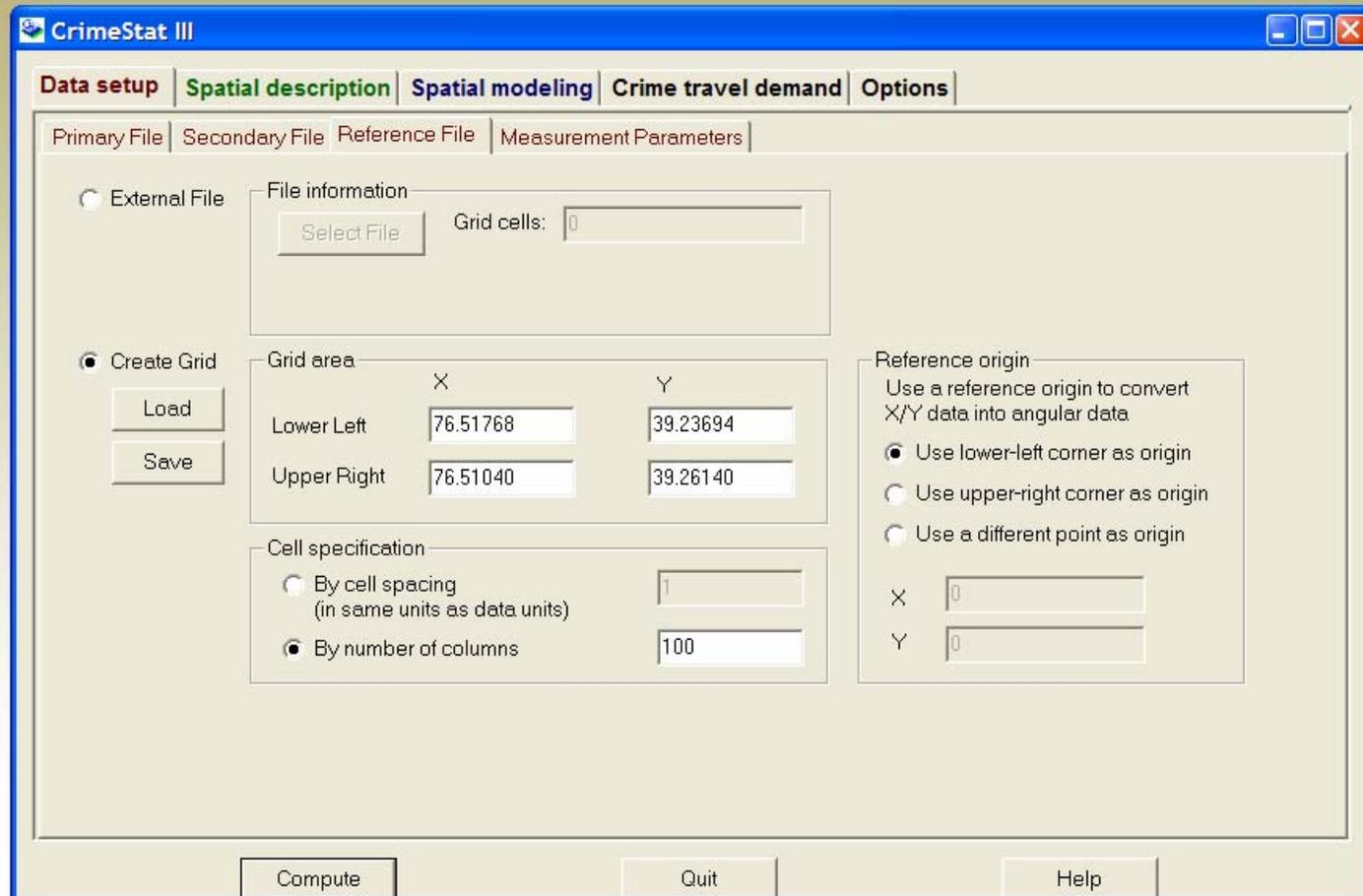
# JTC GP in CrimeStat III

The screenshot displays the 'CrimeStat III' application window. The 'Spatial modeling' tab is active, and the 'Journey-to-Crime' sub-tab is selected. The interface includes several sections for configuring the JTC model:

- Calibrate Journey-to-crime function:** Contains buttons for 'Select data file for calibration', 'Select output file', 'Select kernel parameters', and 'Calibrate!'.
- Journey-to-crime estimation (Jtc):** A checked radio button. Includes an 'Incident file:' dropdown menu set to 'Primary' and a 'Save output to...' button.
- Use already-calibrated distance function:** An unchecked radio button. Includes a text input field, a 'Browse...' button, and a 'Graph' button.
- Use mathematical formula:** A checked radio button. Includes:
  - 'Distribution:' dropdown menu set to 'Truncated negative exponential'.
  - 'Peak likelihood' text input field with value '13.8'.
  - 'Exponent' text input field with value '-0.2'.
  - 'Unit:' dropdown menu set to 'Miles'.
  - 'Peak distance:' text input field with value '0.4'.
  - A text input field with value '1'.
- Draw crime trips:** An unchecked checkbox. Includes a 'Select data file' button and a 'Save output to' button.

At the bottom of the window, there are three buttons: 'Compute', 'Quit', and 'Help'.

# JTC GP in CrimeStat III



The screenshot shows the 'CrimeStat III' application window with the 'Spatial description' tab selected. The window has a blue title bar and standard Windows window controls. Below the title bar are five tabs: 'Data setup', 'Spatial description', 'Spatial modeling', 'Crime travel demand', and 'Options'. Underneath these are four sub-tabs: 'Primary File', 'Secondary File', 'Reference File', and 'Measurement Parameters'. The 'Primary File' sub-tab is active.

The main content area contains the following options and fields:

- External File
- Create Grid

**File information** (under External File):

Select File

**Grid area** (under Create Grid):

|             | X                                     | Y                                     |
|-------------|---------------------------------------|---------------------------------------|
| Lower Left  | <input type="text" value="76.51768"/> | <input type="text" value="39.23694"/> |
| Upper Right | <input type="text" value="76.51040"/> | <input type="text" value="39.26140"/> |

**Cell specification** (under Create Grid):

- By cell spacing (in same units as data units)
- By number of columns

**Reference origin** (under Create Grid):

Use a reference origin to convert X/Y data into angular data

- Use lower-left corner as origin
- Use upper-right corner as origin
- Use a different point as origin

X   
Y

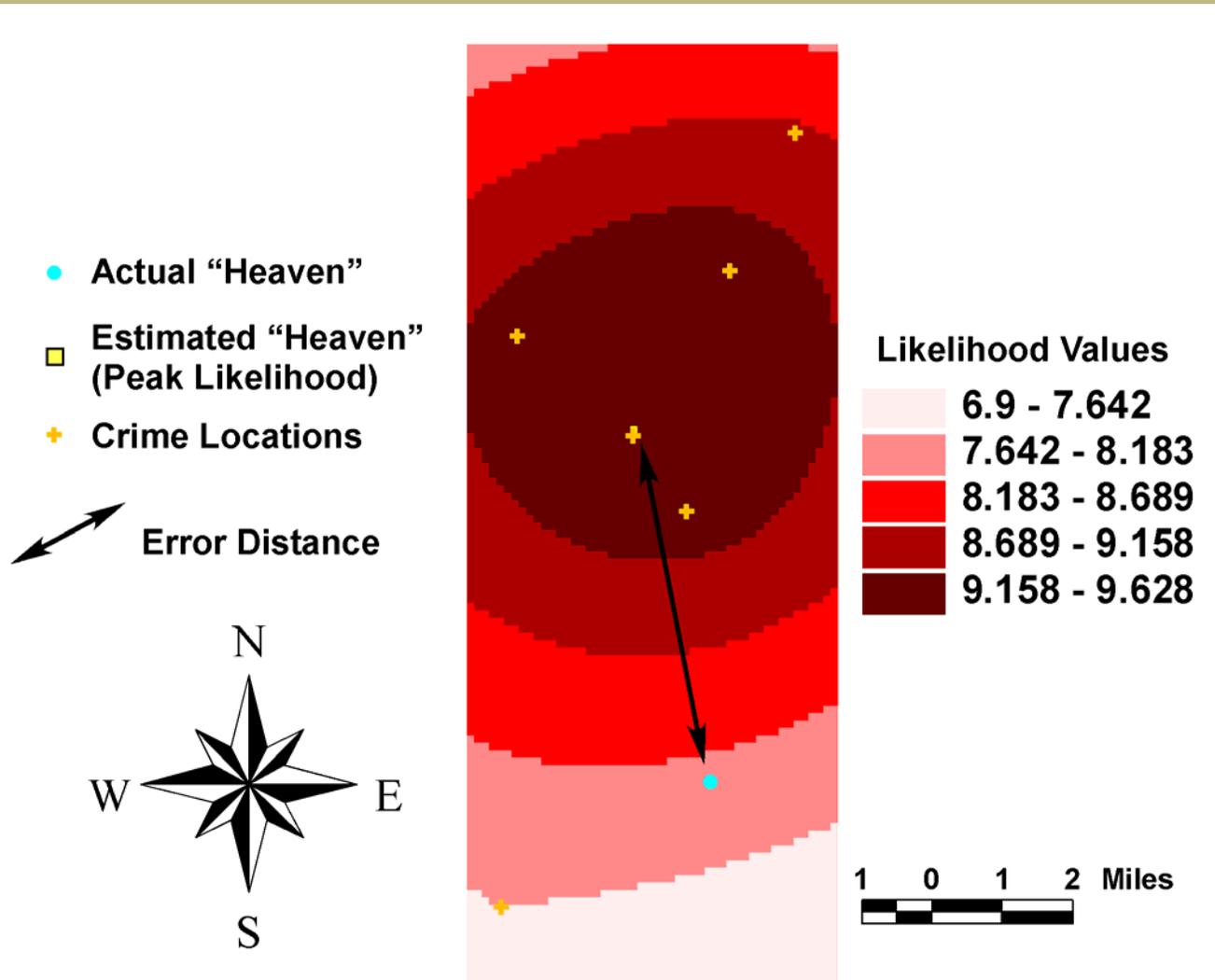
At the bottom of the window are three buttons: 'Compute', 'Quit', and 'Help'.



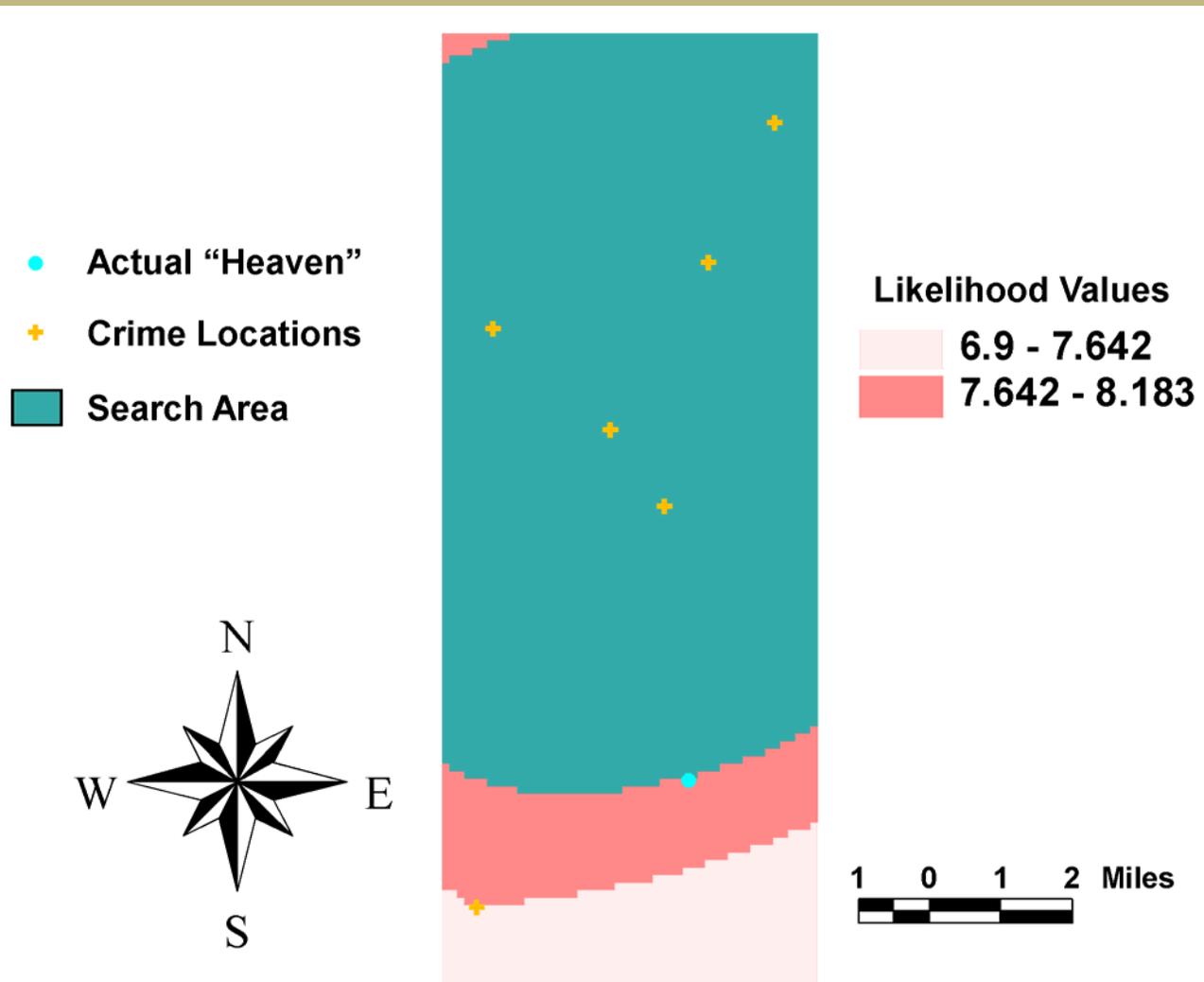
# Measures to compare between GP Results

- Error distance
  - Straight-line distance between the actual and the predicted “haven”
- Search area size
  - Area of all cells with a a probability score equal to or higher than the probability score assigned to the actual haven (hit score)
- Hit score percentage, search cost
  - Ratio between search area and study area
  - The lower the hit score percentage, the more accurate the GP

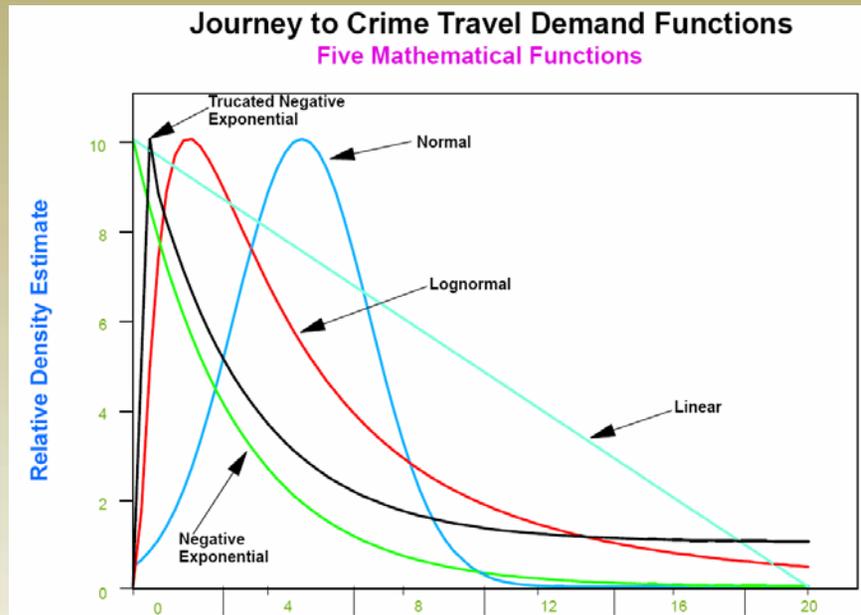
# Comparison Measure - Error Distance



# Comparison Measure – Search Area



# Auto Theft - Linear Function



$$f(d_{ij}) = A + B * d_{ij}$$

- Default values:
- Range of calibrated values:

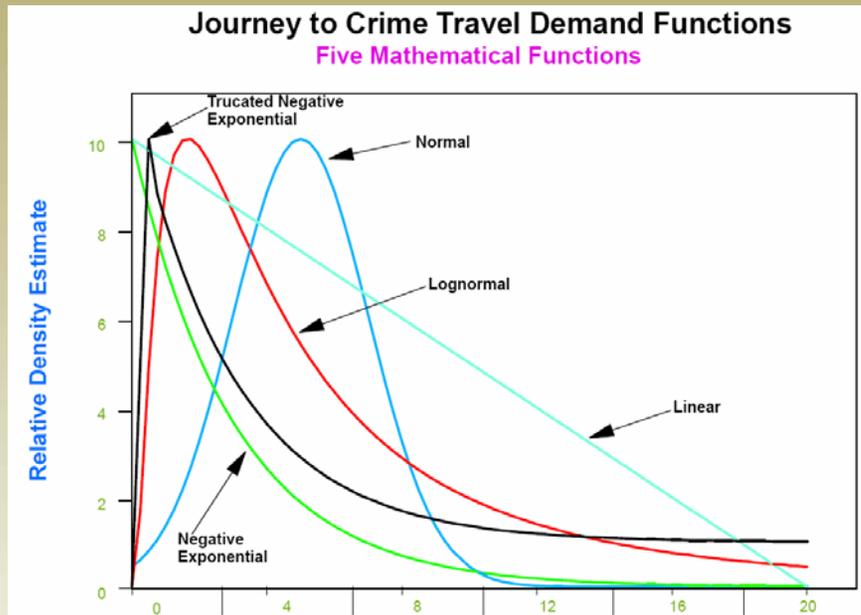
$$A = 1.9, B = -0.06$$

$$A = 2.000 \text{ to } 2.089$$

$$B = -0.0069 \text{ to } -0.0074$$

$$R^2 = 0.302 \text{ to } 0.364$$

# Auto Theft - Negative Exponential Function



$$f(d_{ij}) = A * e^{-B * d_{ij}}$$

- Default values:
- Range of calibrated values:

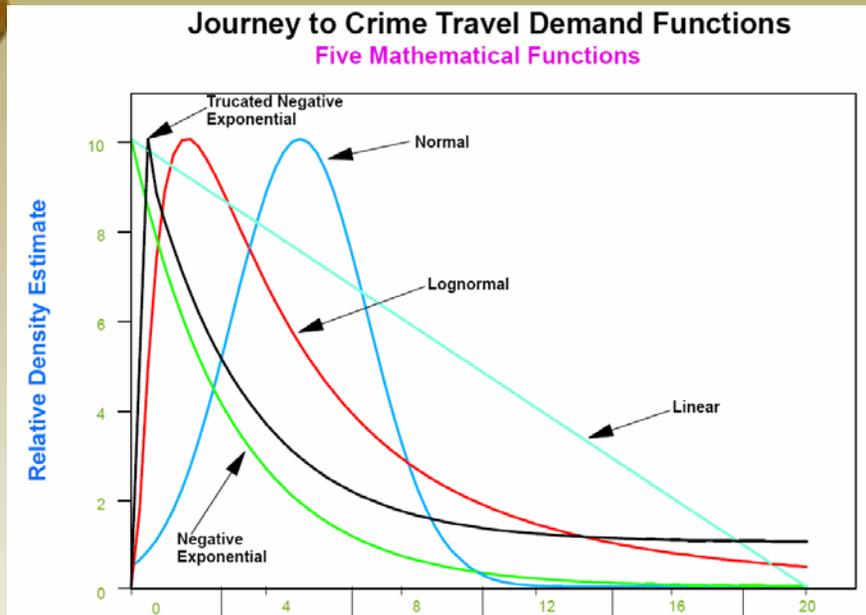
$$A = 1.89, B = -0.06$$

$$A = 0.0069 \text{ to } 0.6480$$

$$B = -0.380 \text{ to } -0.407$$

$$R^2 = 0.399 \text{ to } 0.441$$

# Auto Theft - Normal Function

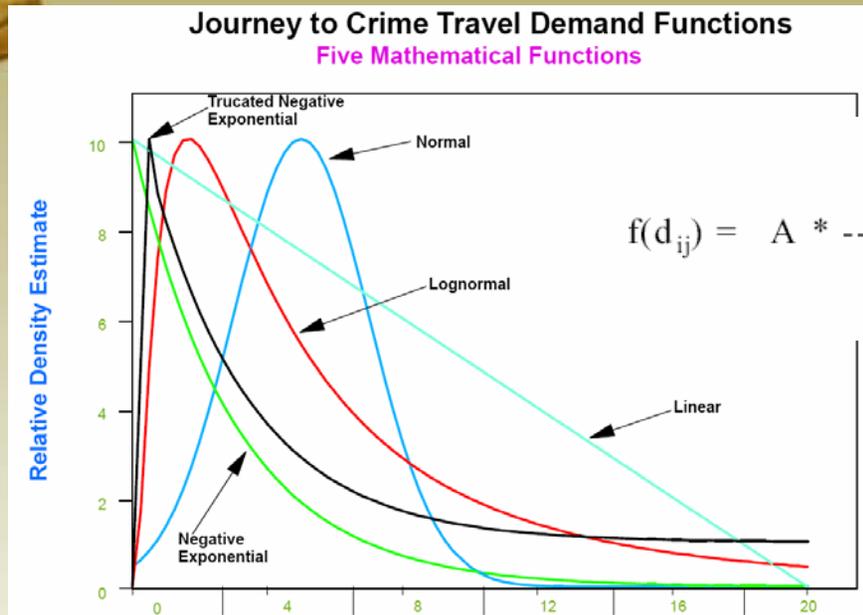


$$Z_{ij} = \frac{(d_{ij} - \text{MeanD})}{S_d}$$

$$f(d_{ij}) = A * \frac{1}{S_d * \text{SQRT}(2\pi)} * e^{-0.5 * Z_{ij}^2}$$

- Default values:      MeanD = 1.9,  $S_d = 4.6$ ,  $A = 29.5$
- Range of calibrated values:      MeanD = 19.7497  
 $S_d = 11.4382$   
 $A = 48.5810$  to  $51.4010$   
 $R^2 = 0.432$  to  $0.500$

# Auto Theft - Lognormal Function

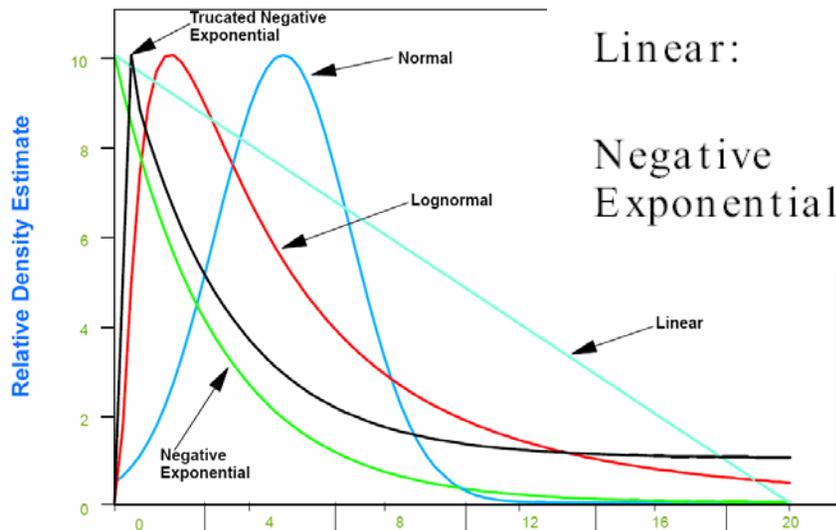


$$f(d_{ij}) = A * \frac{1}{d_{ij}^2 * S_d * \text{SQRT}(2\pi)} * e^{-[\ln(d_{ij}^2) - \text{MeanD}]^2 / 2 * S_d^2}$$

- Default values: MeanD = 4.2, Sd = 4.6, A = 8.6
- Range of calibrated values:
  - MeanD = 19.7497
  - Sd = 11.4382
  - A = -0.3510 to 0.6220
  - R<sup>2</sup> = 0.000

# Auto Theft - Truncated Negative Exponential Function

Journey to Crime Travel Demand Functions  
Five Mathematical Functions



Linear:  $f(d_{ij}) = 0 + B*d_{ij} = B*d_{ij}$  for  $d_{ij} \geq 0, d_{ij} \leq d_p$

Negative Exponential:  $f(d_{ij}) = A*e^{-C*d_{ij}}$  for  $X_i > d_p$

- Default values:

$d_p = 0.4$   
 peak likelihood = 13.8,  
 $C = -0.2$

- Range of calibrated values:

$d_p = 1.375$  to  $2.875$   
 peak likelihood = 5.8394 to 7.6923  
 $C = -0.301$  to  $-0.383$   
 $R^2 = 0.286$  to  $0.395$



## Auto Theft – Comparing Error Distance (m) using a Paired-Samples T Test

| <b>Distance Decay Function</b> | <b>Calibrated</b> | <b>Default</b> | <b>T-Test Statistic</b> | <b>Significance (2-tailed)</b> |
|--------------------------------|-------------------|----------------|-------------------------|--------------------------------|
| Linear                         | 4766              | 4766           |                         |                                |
| Negative Exponential           | 5354              | 5338           | 0.544                   | 0.591                          |
| <b>Normal</b>                  | <b>9545</b>       | <b>4807</b>    | <b>4.916</b>            | <b>0.000</b>                   |
| Lognormal                      | 7181              | 6508           | 1.353                   | 0.186                          |
| Truncated Neg. Exp.            | 5173              | 4994           | 0.375                   | 0.710                          |



## Auto Theft – Comparing Hit Score Percentage using a Paired-Samples T Test

| <b>Distance Decay Function</b> | <b>Calibrated</b> | <b>Default</b> | <b>T-Test Statistic</b> | <b>Significance (2-tailed)</b> |
|--------------------------------|-------------------|----------------|-------------------------|--------------------------------|
| Linear                         | 34.34             | 34.34          |                         |                                |
| Negative Exponential           | 43.29             | 36.31          | 1.772                   | 0.098                          |
| <b>Normal</b>                  | <b>65.85</b>      | <b>43.93</b>   | <b>2.845</b>            | <b>0.013</b>                   |
| Lognormal                      | 51.98             | 39.58          | 1.448                   | 0.17                           |
| Truncated Neg. Exp.            | 43.23             | 46.69          | -0.296                  | 0.771                          |



## Auto Theft – Comparing Search Area (mi<sup>2</sup>) using a Paired-Samples T Test

| <b>Distance Decay Function</b> | <b>Calibrated</b> | <b>Default</b> | <b>T-Test Statistic</b> | <b>Significance (2-tailed)</b> |
|--------------------------------|-------------------|----------------|-------------------------|--------------------------------|
| Linear                         | 34.34             | 34.34          |                         |                                |
| Negative Exponential           | 40.55             | 30.08          | 2.056                   | 0.059                          |
| <b>Normal</b>                  | <b>62.06</b>      | <b>30.87</b>   | <b>3.189</b>            | <b>0.007</b>                   |
| <b>Lognormal</b>               | <b>62.04</b>      | <b>30.86</b>   | <b>3.189</b>            | <b>0.007</b>                   |
| Truncated Neg. Exp.            | 32.16             | 36.98          | -1.121                  | 0.281                          |



## Summary of Results

- For serial offenses involving Auto Thefts, comparison measurements to not show significant different results, whether or not distance decay functions are calibrated.
- Unexpectedly, the normal distance decay function using the default parameters produces more accurate GP for all three comparison measurements.
- Preliminary results indicate that spending time and resources to calibrate distance decay functions individually may not be necessary.



## Future Research

- Redo this analysis with marauder type serial offenders only (i.e., remove the commuter type serial offenders).
- Redo this comparison analysis with a Bayesian JTC routine.
- As always, redo this analysis for other study areas and different set of serial offense data.